Is careless responding also a problem in face-to-face mode? Analysis of PIAAC noncognitive data.

Marek Muszyński
Tomasz Żółtak
Artur Pokropek

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Response biases in self-reports

- Self-report – one of the fundamental research methods in social sciences (Paulhus & Vazire, 2007)

- only method in 30% of papers (Woszczynski & Wittman, 2004); present in more than 60% of papers (Brutus et al., 2010)

- ca. 55% papers from *American Sociological Review* and ca. 39% from *American Journal of Sociology* used self-reports (Bruckner, 2009)

- Important part of every large-scale assessment project (PIAAC, PISA, TIMSS, etc.)

- Response biases danger (Khorramdel et al., 2017, 2019; Lechner et al., 2019; Palczyńska & Rynko, 2021; Rammstedt et al., 2017; Ulitzsch et al., 2021)

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Response biases in self-reports

Error sources

- Respondent
- Interviewer
- Instrument
  - ...

Response biases

- Careless responding
- Response styles
- Response sets

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Careless responding

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Careless responding

- Careless/insufficient effort responding (C/IER)
- Responding without sufficient effort, understanding or regard towards items’ content and/or survey instructions (Meade & Craig, 2012; Huang et al., 2012, 2015; Ullitzsch et al., 2021)
- Satisficing (Krosnick, 1991)
- Due to low interest, fatigue, no social contract, environmental/psychological distractions (Meade & Craig, 2012)
- May follow different patterns: (pseudo)random responding, straightlining, fixed lining (diagonal, snake), etc.
Careless responding

- Careless/insufficient effort responding (C/IER) does not reflect traits to be measured (Ulitzsch et al., 2021)

- Threat to construct and criterion-related validity (Haung et al., 2015; McGrath et al., 2010)

- ... and to scales’ psychometric properties, e.g. reliability, factor structure (De Simone et al., 2018)
Careless responding – why?

Careless responding

- Previous studies concentrate on self-completion mode and on web-surveys
- Less evidence on C/IER in the CAPI-mode assessments

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
PIAAC – used variables

- Data from:
  - 209,997 participants
  - 37 countries from all three PIAAC cycles
  - C/IER indices for “H” part of the questionnaire: reading, writing, numeracy, and ICT skills in everyday life (25 items, 1:5)
  - 2,256 participants eliminated due to all missings for the “H” part (1.07% of sample)

Careless responding – how?
Careless responding – how?

C/IER indices used

- “careless” R package (Yentes & Wilhelm, 2018/2021)
- Mahalanobis distance
- Longstring (Meade & Craig, 2012; Kim et al., 2019)

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
C/IER indices cut-off thresholds

- Liberal: 95th quantile Mahalanobis and IRV, 75% longstring
- Conservative: 99th quantile Mahalanobis and IRV, 100% longstring

(Curran, 2016; Ulitzsch et al., 2021)
Indicators of C/IER clearance

- Internal consistency (Cronbach’s Alpha)
- Construct validity – correlation with other background scales
- Criterion-related validity – correlation with cognitive tests
### Results – internal consistency

<table>
<thead>
<tr>
<th>Scale</th>
<th>Overall</th>
<th>C/IER liberal</th>
<th>C/IER conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>0.73 (8 items)</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>Writing</td>
<td>0.50 (4 items)</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>Numeracy</td>
<td>0.73 (6 items)</td>
<td>0.72</td>
<td>0.73</td>
</tr>
<tr>
<td>ICT use</td>
<td>0.58 (7 items)</td>
<td>0.58</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Careless responding - results

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Results – construct validity

<table>
<thead>
<tr>
<th>Scale</th>
<th>Overall</th>
<th>C/IER liberal</th>
<th>C/IER conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>0.51</td>
<td>0.52</td>
<td>0.51</td>
</tr>
<tr>
<td>Writing</td>
<td>0.38</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>Numeracy</td>
<td>0.43</td>
<td>0.44</td>
<td>0.43</td>
</tr>
<tr>
<td>ICT use</td>
<td>0.35</td>
<td>0.36</td>
<td>0.35</td>
</tr>
</tbody>
</table>
### Results – criterion-related validity

<table>
<thead>
<tr>
<th>Scale</th>
<th>Overall</th>
<th>C/IER liberal</th>
<th>C/IER conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>0.43</td>
<td>0.45</td>
<td>0.44</td>
</tr>
<tr>
<td>Writing</td>
<td>0.35</td>
<td>0.38</td>
<td>0.36</td>
</tr>
<tr>
<td>Numeracy</td>
<td>0.35</td>
<td>0.40</td>
<td>0.38</td>
</tr>
<tr>
<td>ICT use</td>
<td>0.34</td>
<td>0.35</td>
<td>0.34</td>
</tr>
</tbody>
</table>
Careless responding - conclusions

Conclusions

- C/IER seems to be a limited threat to PIAAC data...
- ... or maybe the thresholds were too liberal as...
- ... liberal thresholds enable to increase validity (though only a tiny bit)
- Spurious reliability due to (alleged) straightlining (+ no reversed items)
- Empirical mode comparisons seem to be rare in the field?
- Experiments in CAPI surveys are also limited
- Research should continue as LSAs are not absolved from response biases (e.g. Goldhammer et al., 2017; Lechner et al., 2019; Rammstedt et al., 2017)
Future studies/ideas

- Simulation studies to test sensitivity of various questionnaire data to C/IER in LSA contexts
- Interviewer-level analysis (Menold & Kemper, 2013):
  - Observational data (Loosveldt & Beullens, 2017)
  - Interviewer characteristics (Bittman, 2021; Jacobs et al., 2019)
- If CAWI used – paradata analysis to enhance classic indices (Horwitz et al., 2017, 2019; Ulitzsch et al., 2021)
- More advanced C/IER indices (Mansolf & Reise, 2018; Yentes, 2020; also Nissen et al., 2016)
- Non-threshold approach (PCA; Huang et al., 2015; LCA; Meade & Craig, 2012; Ulitzsch et al., 2021)

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)
Future studies/ideas

- C/IER in cognitive assessment vs. in background questionnaire
- % of sample screened out as C/IER
- Interviewer effects
- C/IER as a trait – use panel data (Canadian PIAAC?)
- Cultural differences:
  - Mechanisms
  - Differences: perform analysis on different ILSAs (PIAAC, PISA, TIMSS, ESS, etc.), and on different waves (e.g. all PISA waves)
Thank you!

rstyles.ifispan.edu.pl

www.researchgate.net/lab/Artur-Pokropek-Lab

marek.muszynski@ifispan.edu.pl

This project is financed by the National Science Centre (NCN) research grant (2019/33/B/HS6/00937)